

Adaptation française de la section III du Test for Auditory Comprehension of Language – Revised et développement de normes régionales

French Adaptation of Section III of the Test for Auditory Comprehension of Language – Revised, and Development of Regional Standards

par • by

Nicole Alma, CCO
Conseil des écoles catholiques
de Halton
Burlington, Ontario

Sylvie Gautheron-Stone,
CCO
Hôpital général de North York
North York, Ontario

Lynne Latulippe, MHSc
Conseil des écoles séparées
catholiques romaines
de Dufferin et Peel
Mississauga, Ontario

Linda Siegel, PhD
Université de Colombie-
Britannique
Vancouver, Colombie-
Britannique

ABRÉGÉ

Le but de cette étude était le développement d'un instrument d'évaluation français de la compréhension syntaxique d'élèves du jardin à la sixième année (5 ans à 11 ans) fréquentant des écoles de langue française dans le centre-sud de l'Ontario et vivant en milieu linguistique minoritaire. Pour y parvenir, la section III du test anglais *Test for Auditory Comprehension of Language – Revised* (Carrow-Woolfolk, 1985) a été traduite et adaptée en français puis a été soumise individuellement à 977 élèves après obtention de renseignements portant sur leur vécu linguistique et leur rendement académique. Il existe une différence significative entre la moyenne des scores des élèves plus exposés au français et celle des élèves moins exposés au français jusqu'en troisième année (huit ans) où commence un effet plafond. Les analyses montrent que ce test est fidèle et valide. Les normes régionales développées se sont avérées différentes des normes du test original.

ABSTRACT

The aim of this study was to develop a French-language assessment tool for syntactic comprehension in children from kindergarten to grade six (from 5 to 11 years) attending French-language schools in south-central Ontario and residing in a minority-language environment. To this end, Section III of the English-language *Test for Auditory Comprehension of Language – Revised* (Carrow-Woolfolk, 1985) was translated and adapted into French. It was then administered individually to 977 students after obtaining information on their linguistic experience and their performance at school. There was a significant difference between the average score of students more exposed to French versus those of students less exposed until the third grade (8 years), where a "ceiling effect" begins. Analyses have shown this test to be consistent and valid. The regional standards that were developed were different from those of the original test.

MOTS CLÉ

langage réceptif • compréhension syntaxique • *TACL-R* • normes franco-ontariennes • indice de francité • langue minoritaire

Parmi les problèmes associés à l'utilisation d'un test figure celui de l'usage des normes. Peut-on se servir des normes d'un test lorsque ce dernier a été standardisé sur une population dont la réalité linguistique est différente de la population qui est évaluée? Telle est la question à laquelle sont confrontés les orthophonistes travaillant dans les écoles de langue française dans le centre-sud de l'Ontario.

En Ontario, les francophones ne représentent que 5.4% de la population totale de la province (Statistique Canada, 1992). Dans la région centre-sud où l'on trouve surtout des centres urbains, cette population francophone est d'origine franco-ontarienne, acadienne, québécoise et comprend également des nouveaux immigrants (Heller & Barker, 1988; Heller, Lévy, Mougéon, & Quirouette, 1986). Sur le plan social, culturel et économique, ces communautés francophones sont hétérogènes

(Mougeon & Heller, 1986). De plus, étant minoritaire, la population francophone du centre-sud de l'Ontario est fortement influencée par l'anglais (Latulippe, 1992). Dans cette région, les mariages mixtes entre francophones et non-francophones sont communs (Mougeon, 1979). Par exemple, d'après Heller et al. (1986) dans la région torontoise, les mariages mixtes représentent 40% des mariages de la communauté francophone. Tenant compte de ces éléments, il n'est pas surprenant que dans les écoles francophones, nous retrouvions des élèves avec différents niveaux de francité. Nous avons emprunté à Desjarlais (1980) cette terminologie qui se réfère au niveau de contact avec la langue dans le contexte familial, social et culturel.

Plusieurs enquêtes dont celle de Desjarlais (1980) ont révélé que, afin de se conformer à la loi sur l'Éducation de l'Ontario et à la Charte canadienne des droits et libertés, les écoles de langue française acceptent non seulement des élèves qui communiquent en français au foyer mais également des élèves qui communiquent régulièrement en anglais ou dans une autre langue au foyer. Mougeon, Heller, Béniak, et Canale (1984) précisent que dans les écoles de langue française des milieux minoritaires, il existe trois groupes d'élèves soit franco-dominant, soit bilingue, soit anglo-dominant. Hébrard et Mougeon (1975) ont étudié la question de l'effet de la langue (ou des langues) parlée(s) dans les foyers franco-ontariens sur le comportement linguistique des enfants aux paliers élémentaire et secondaire. Ils indiquent que "la langue parlée le plus souvent entre les parents et les enfants est un facteur crucial dans l'acquisition linguistique de l'enfant dans un milieu bilingue" (p. 68).

Dans le milieu scolaire, les orthophonistes sont donc confrontés à des niveaux de compétence linguistique variés en français. Par ailleurs, la plupart des tests de langage disponibles ne sont que des traductions-maison de tests anglais, traductions faites individuellement et non standardisées auprès de cette population-cible. De nombreux orthophonistes ont rejeté l'utilisation des tests français de France car ils ne reflètent pas la réalité linguistique et culturelle de l'Ontario. Les orthophonistes oeuvrant dans d'autres milieux francophones au Canada sont confrontés à ce même problème.

Il existe une batterie de tests provenant du Québec, les *Tests de langage dudley-delage* (Dudley & Delage, 1980) dont les normes ont été développées sur une population québécoise. L'utilisation de ces tests en Ontario ne nous donne malheureusement que des informations limitées. D'une part, le contexte linguistique est différent et d'autre part ce test n'a été standardisé qu'au niveau du jardin à la deuxième année. Un autre outil standardisé s'adressant aux populations canadiennes-françaises, l'*Échelle de vocabulaire en images Peabody* (Dunn, Thériault-Whalen, & Dunn, 1993), vient d'être mis sur le marché. Cette adaptation récente du *Peabody Picture Vocabulary Test — Revised* (Dunn & Dunn, 1991) a été standardisée auprès

de populations francophones dans les provinces du Québec, de l'Ontario, de l'Alberta et du Nouveau-Brunswick. Ce test explore la compréhension lexicale qui n'est qu'un des nombreux aspects des habiletés langagières (Rondal & Seron, 1985).

Au niveau de l'évaluation langagière des élèves bilingues (quelles que soient les langues parlées), Roussel (1991) présente dans une bibliographie descriptive une soixantaine de tests de communication disponibles pour l'évaluation de ces élèves. Ces tests représentent une gamme d'outils pour évaluer des élèves dont la première langue est autre que l'anglais (e.g., le chinois, l'espagnol et l'hindi). Bien que ces tests soient disponibles, leur efficacité demeure limitée car peu présentent des preuves de validité et de fidélité et seul un nombre restreint d'entre eux est standardisé. De plus, une recherche menée par Wilcox et McGuinn-Aasby (1988) sur une traduction espagnole du *Test for Auditory Comprehension of Language* (Carrow, 1973) a bien mis en évidence qu'il était nécessaire d'obtenir des normes différentes si le test utilisé était un test traduit.

Peu de recherches se sont penchées sur la question de la création, de l'adaptation et de la standardisation d'outils d'évaluation desservant les populations bilingues (français-anglais) vivant en milieu francophone minoritaire. Une revue de la littérature n'a décelé aucun article ou aucune étude portant sur les aptitudes langagières réceptives au niveau de la syntaxe dans notre population-cible.

Cet article relate un projet de recherche dont le but était de développer un outil d'évaluation en français valide et fiable permettant de mesurer les habiletés de langage réceptif au niveau syntaxique des élèves du palier élémentaire fréquentant les écoles de langue française en Ontario. Afin d'obtenir cet instrument d'évaluation, deux étapes se sont avérées nécessaires : tout d'abord nous avons procédé à la traduction et à l'adaptation de la section III (*Elaborated Sentences*) du *Test for Auditory Comprehension of Language — Revised (TACL-R)* de Carrow-Woolfolk (1985) et ensuite nous avons élaboré des normes régionales représentant les habiletés langagières syntaxiques réceptives de la population qui nous intéresse, c'est-à-dire les élèves du jardin à la sixième année des écoles de langue française du centre-sud de l'Ontario. Il est à noter qu'en Ontario le terme jardin désigne l'année de scolarisation qui précède la première année. Dix hypothèses ont servi de fil conducteur à cette recherche :

1. Les scores au test augmenteront au fur et à mesure que l'âge des élèves augmente.
2. Les scores au test augmenteront au fur et à mesure que le niveau scolaire augmente.
3. Les élèves ayant un indice de francité élevé réussiront mieux au test que les élèves ayant un indice de francité faible.
4. Plus le rendement académique sera élevé, plus les résultats au test seront élevés.

5. Les élèves qui reçoivent de l'aide académique supplémentaire réussiront moins bien au test que les élèves qui ne reçoivent pas cette aide.

6. Les élèves ayant des difficultés de langage réceptif auront des scores moins élevés que les élèves sans difficultés de langage.

7. Les normes développées lors de ce projet seront différentes des normes du test original.

8. L'ordre de difficulté des items sera semblable à celui retrouvé dans le test original.

9. Les validités de construit, de contenu et celle liée à un critère seront établies.

10. La fidélité du test sera démontrée.

Méthodologie

Traduction du test

Standardisé auprès d'une population américaine de 1003 enfants âgés de trois à neuf ans, le *TACL-R* est un test de langage réceptif. C'est un test valide et reconnu qui permet d'évaluer efficacement les enfants ayant un retard de langage (Bankson, 1989; Haynes, 1989).

Le *TACL-R* est divisé en trois sections de 40 items (I, *Word Classes and Relations*; II, *Grammatical Morphemes*; et III, *Elaborated Sentences*). Les sections I et II n'ont pas été retenues pour ce projet puisque la section I n'examine que le vocabulaire réceptif et la section II ne se prête pas à une traduction en français à cause des différences linguistiques. La troisième section sur laquelle porte ce projet comprend 40 planches illustrant des structures syntaxiques classées par ordre de difficulté croissante. Sur chaque planche, il y a trois images en noir et blanc. Une des images correspond à la phrase stimulus.

La traduction de la section III qui a été faite par trois orthophonistes ayant une formation en linguistique a nécessité des adaptations pour respecter les différences linguistiques entre le français et l'anglais. Chaque item a été analysé en détail au niveau syntaxique, au niveau lexical et en fonction des pièges créés par les images. Puis, chaque item a été traduit et revu afin d'obtenir, autant que possible, une traduction représentant les mêmes formes lexicales et syntaxiques que dans l'item original.

La traduction de l'item 24 "Besides the baseball glove, she bought a record." était réalisable. Néanmoins l'item n'a pas été retenu dans la version française car la traduction ("En plus du gant de baseball, elle a acheté un disque.") supprimait l'ambiguïté suggérée par les images.

Sur les images 9, 19, 20 et 25 les mots anglais ont été traduits ou supprimés. La traduction de certaines des phrases stimuli a donné lieu à des discussions quant au choix du vocabulaire. Par exemple le mot "lady" a été traduit par "dame" plutôt que par "femme" car dans le contexte linguistique franco-ontarien "dame" est plus souvent

employé par les enfants. Le raisonnement a été le même pour le choix de "monsieur" comme traduction de "man" plutôt que "homme".

Choix de la population

La population cible de ce projet est constituée d'élèves fréquentant des écoles de langue française dans des conseils scolaires de la région centre-sud de l'Ontario. Les quatre conseils scolaires qui ont participé au projet sont les Conseils catholiques de Halton, de Dufferin et Peel, de Hamilton-Wentworth et le Conseil des écoles françaises de la communauté urbaine de Toronto (CEFCUT).

L'échantillon

Les élèves du jardin à la sixième de quatre écoles, soit une dans chaque conseil scolaire, ont participé à ce projet. Ces écoles ont été choisies selon trois critères : que l'école ait une bonne représentation des divers niveaux socio-économiques des élèves du conseil, que la direction de l'école soit intéressée à participer au projet et qu'un local adéquat soit disponible pour permettre l'évaluation des élèves. Dans le test original, les normes ont été établies à partir du jardin à la quatrième année. Pour la recherche, les normes ont été établies à partir du jardin et nous avons évalué les élèves jusqu'en sixième année car dans la version originale du *TACL-R* plusieurs items n'avaient pas été réussis par 90% des élèves en quatrième année. De plus, le français étant une langue minoritaire dans le centre-sud de l'Ontario, il est possible que l'acquisition de certaines structures prenne plus de temps que dans un milieu unilingue. Nous avons donc décidé de faire participer les élèves jusqu'en sixième année. Parmi les 1086 élèves fréquentant les quatre écoles choisies, 977 (soit 90%) ont accepté de participer à l'étude.

Collecte des données

Formulaire de consentement et questionnaire destinés aux parents. On a décidé de distribuer à tous les parents des élèves du jardin à la sixième année une lettre décrivant les objectifs du projet ainsi que deux documents à remplir : un formulaire d'autorisation (ou de refus) de participation au projet et un court questionnaire dont le but était de recueillir des informations sur la langue parlée le plus souvent par l'enfant selon les personnes auxquelles il s'adresse (membres de sa famille, gardienne et amis). Le parent devait indiquer un seul choix de langue par interlocuteur.

Questionnaire destiné aux enseignants et enseignantes. Les enseignants et enseignantes des élèves qui ont participé à l'étude ont rempli un court questionnaire. Ce questionnaire avait comme objectif d'obtenir des informations sur la langue dominante de l'élève, son rendement académique (faible, moyen, fort), et de déterminer s'il ou elle recevait de l'aide supplémen-

taire en raison de difficultés académiques. Le rendement académique de l'élève dans son travail journalier a été coté fort, moyen ou faible dans les cinq domaines suivants : mathématiques, lecture, rédaction, expression orale et compréhension orale.

Identification des élèves ayant des difficultés de langage. Les orthophonistes des quatre conseils scolaires ont préparé une liste des élèves identifiés comme ayant des difficultés de langage.

Passation du test. Après avoir reçu le consentement de la part des parents, deux agentes de recherche ont évalué les élèves. Chaque élève a été rencontré individuellement dans un local approprié à la passation du test.

Analyse statistique des données

Suite à la collecte des données, une base de données informatisée a été créée en utilisant le logiciel Statsview version MacIntosh. L'analyse a été faite au moyen de ce même logiciel et ainsi que du logiciel SPSS version MacIntosh.

Critères de dominance langagière

A partir des réponses faites par les parents et le personnel enseignant aux questionnaires, un indice de francité a été élaboré pour identifier les élèves ayant un haut niveau de contact avec le français et ceux ayant un contact moindre avec le français.

L'objectif était que cet indice reflète l'importance du français dans le vécu langagier de l'élève. Afin de pondérer les résultats, un nombre de points plus important a été accordé lorsque le français était parlé avec la mère ainsi qu'avec les frères et soeurs, ces deux situations étant un indice d'une forte influence du

Tableau 1. Attribution des points pour le calcul de l'indice de francité.

	fran	fran et ang	fran et aut	anglais	autre
Langue parlée par l'enfant avec :					
mère	4	2	2	0	0
père	2	1	1	0	0
frères/soeurs	4	2	2	0	0
amis	2	1	1	0	0
autre personne	2	1	1	0	0
langue parlée entre les parents	2	1	1	0	0
langue dominante d'après l'enseignant(e)	2	1	1	0	0
total possible:	18				

français dans la vie de l'élève. Lorsque, malgré les directives, les répondants et les répondantes au questionnaire indiquaient deux choix de langue (e.g., français et anglais, français et autre), le nombre de points était divisé par deux.

Le seuil de francité a été établi à 11/18, soit 0.6. Un résultat supérieur ou égal à 0.6 est signe d'un indice de francité élevé alors qu'un résultat inférieur à 0.6 correspond à un indice de francité faible.

Résultats et discussion

Hypothèse 1 : les scores au test augmenteront au fur et à mesure que l'âge des élèves augmente

L'hypothèse 1 a été confirmée par les résultats présentés dans le Tableau 2. De plus, une ANOVA démontre que la différence est significative [$F(7,969) = 162.38; p < .0001$].

Tableau 2. Statistiques descriptives des résultats au test selon l'âge.

groupe d'âge	n	moyenne des réponses correctes	écart-type	erreur-type	étendue
5;0 - 5;11	119	21.01	4.77	0.44	10 - 32
6;0 - 6;11	139	25.55	4.35	0.37	16 - 35
7;0 - 7;11	133	29.33	3.59	0.31	19 - 35
8;0 - 8;11	162	30.53	3.89	0.31	14 - 38
9;0 - 9;11	139	31.63	4.11	0.35	15 - 39
10;0 - 10;11	136	32.99	3.18	0.27	20 - 39
11+	145	34.27	2.98	0.25	19 - 39

Hypothèse 2 : les scores au test augmenteront au fur et à mesure que le niveau scolaire des élèves augmente

Il a été confirmé que les moyennes des scores au test augmentent selon le niveau scolaire et une analyse de la variance indique que la différence trouvée est significative [$F(6,970) = 194.02; p < .0001$].

Tableau 3. Statistiques descriptives des résultats au test selon le niveau scolaire.

niveau scolaire	n	moyenne des réponses correctes	écart-type	erreur-type	étendue
jardin	119	20.98	4.88	0.45	10 - 32
1 ^{ère}	135	25.17	4.39	0.38	14 - 35
2 ^e	148	29.06	3.88	0.32	14 - 36
3 ^e	168	30.55	3.82	0.30	16 - 38
4 ^e	127	31.98	3.89	0.35	15 - 39
5 ^e	139	32.95	3.12	0.26	19 - 38
6 ^e	141	34.40	2.73	0.23	28 - 39

Tableau 4. Répartition des élèves par indice de francité.

niveau scolaire	nombre total d'élèves	indice faible		indice élevé	
		n	%	n	%
jardin	119	90	75.6	29	24.4
1 ^{ère}	135	99	73.3	36	26.7
2 ^e	148	112	75.6	36	24.4
3 ^e	168	123	73.2	45	26.8
4 ^e	127	89	70.1	38	29.9
5 ^e	139	106	76.2	33	23.8
6 ^e	141	97	68.7	44	31.3
total	977	716	73.3	261	26.7

Tableau 5. Statistiques descriptives des résultats au test selon l'âge et l'indice de francité.

groupe d'âge/indice de francité	n	moyenne des réponses correctes	écart-type	erreur-type	étendue
5;0 - 5;11					
faible	90	20.34	4.71	0.50	11 - 32
élevé	29	23.07	4.45	0.83	10 - 31
6;0 - 6;11					
faible	102	24.92	4.32	0.43	16 - 33
élevé	37	27.30	4.01	0.66	19 - 35
7;0 - 7;11					
faible	99	28.84	3.49	0.35	19 - 35
élevé	34	30.77	3.53	0.61	19 - 35
8;0 - 8;11					
faible	126	30.42	3.72	0.33	16 - 38
élevé	36	30.92	4.49	0.75	14 - 38
9;0 - 9;11					
faible	93	31.46	3.94	0.41	20 - 38
élevé	46	31.96	4.45	0.66	15 - 39
10;0 - 10;11					
faible	100	32.77	3.01	0.30	26 - 38
élevé	36	33.59	3.59	0.60	20 - 39
11+					
faible	102	34.17	3.14	0.31	19 - 39
élevé	43	34.51	2.58	0.39	28 - 38

Hypothèse 3 : les élèves ayant un indice de francité élevé réussissent mieux au test que les élèves ayant un indice de francité faible

Le Tableau 4 indique qu'en dehors du contexte scolaire, près de trois quart des élèves de l'échantillon ont un vécu langagier dans lequel le français n'a pas la place prépondérante.

Le Tableau 5 présente les résultats des élèves en tenant compte de leur indice de francité.

Un examen visuel suggère que les élèves avec un indice de francité élevé ont mieux réussi que ceux ayant un indice de francité faible. Cependant cette différence diminue à partir de la troisième année (vers l'âge de huit ans). Des analyses de la variance effectuées selon le niveau scolaire et le groupe d'âge, comparant les résultats des élèves avec un indice de francité élevé à ceux des élèves avec un indice de francité faible, démontrent qu'il existe des différences significatives entre ces résultats jusqu'en deuxième année (et jusqu'à l'âge de 7;11 ans) ($p < .05$).

Une analyse de la variance à deux critères tenant compte de l'indice de francité [$F(1) = 23.62; p < .001$] et de l'âge [$F(6) = 131.28; p < .001$] a démontré que, bien que ces deux critères soient statistiquement significatifs, il n'y a pas d'interaction entre eux [$F(6) = 1.66; p < .13$].

Une analyse semblable effectuée en tenant compte des niveaux scolaires [$F(6) = 137.4; p < .001$] et non plus de l'âge donne le même résultat. [ANOVA indice de francité X niveau scolaire $F(6) = 1.85; p < .087$]. Le Tableau 2 montre l'augmentation des scores d'un âge à l'autre. Cependant, l'augmentation des scores est de plus en plus faible à partir de l'âge de huit ans (ou de la troisième année). D'autre part, l'aplatissement entre les scores obtenus par les élèves à francité élevée et ceux obtenus par les élèves à francité faible suggère qu'il faut interpréter avec précaution l'absence de différences significatives entre les résultats des deux groupes à partir de la troisième année. Ceci signifie que jusqu'en deuxième année, le test est sensible pour mesurer le développement des acquis dans le domaine de la syntaxe réceptive. Ces résultats nous permettent également d'émettre l'hypothèse que les acquis syntaxiques mesurés par le test (et ceux-ci uniquement) sont semblables à partir de la troisième année. Cependant, ceci ne permet pas de conclure que les élèves ayant un indice de francité faible et ceux ayant un indice de francité élevé ont le même niveau de compréhension syntaxique. Par exemple, on pourrait retrouver des différences entre ces groupes quant à la compréhension de structures syntaxiques plus complexes que celles évaluées par ce test.

Élèves exposés à une troisième langue. Parmi les sujets évalués lors de ce projet de recherche, 170 ont été identifiés comme étant exposés à une troisième langue. Un total de 27 langues autres que le français a été recensé dans le questionnaire des parents. Le persan, l'arabe et le somalien sont les trois langues qui reviennent le plus souvent.

En appliquant à la troisième langue le même système de points que celui utilisé pour déterminer l'indice de francité des élèves, nous avons détecté 40 cas dont le niveau de contact avec cette troisième langue est élevé. Étant donné que ces 40 élèves sont répartis dans différents niveaux scolaires et différents groupes d'âge, il est impossible de comparer statistiquement les résultats de ce groupe exposé à une troisième langue à ceux des autres groupes. Cependant, l'utilisation des tableaux de normes révèle que presque un quart (soit neuf élèves) de ces 40 élèves se situe sous le quinzième rang centile. D'après les questionnaires, il ressort que ces neuf élèves éprouvent des difficultés académiques et que plusieurs reçoivent de l'aide supplémentaire en francisation. Une vérification auprès des professionnels et professionnelles oeuvrant auprès de ces neuf élèves indique que toutes et tous vivent un processus de transculturation.

Hypothèse 4 : plus le rendement académique est élevé, plus les résultats au test seront élevés

Les enseignants et enseignantes des élèves qui ont participé à l'étude ont rempli un questionnaire pour chaque élève indiquant son rendement académique (fort, moyen ou faible) dans les domaines suivants : compréhension orale, expression orale, lecture, rédaction et mathématiques.

Un examen visuel des réponses faites par les enseignants et les enseignantes montre que certains d'entre eux ont jugé que la majorité de leurs élèves avait un rendement académique fort, qu'un groupe plus restreint avait un rendement dans la moyenne et que seuls quelques élèves étaient classés faibles. Ces résultats ressemblent à ceux obtenus par Egan et Archer "[les enseignants et les enseignantes] were reluctant to use the extreme categories. In addition they overused the upper quintiles" (1985, p. 32). Étant donné le peu d'élèves ayant un indice de francité élevé, il n'a pas été possible de vérifier statistiquement l'existence d'une différence significative entre les moyennes des groupes faible, moyen et fort. L'analyse a donc été effectuée sur l'ensemble des élèves sans tenir compte de l'indice de francité. Considérons le rendement académique en compréhension orale déterminé par le questionnaire des enseignants et des enseignantes. Les résultats des ANOVAs du rendement académique selon le niveau scolaire montrent l'existence de différences significatives.

Tableau 6. Statistiques descriptives du rendement académique en compréhension orale par niveau et résultats des ANOVAs.

niveau scolaire		<i>n</i>	moyenne des réponses correctes	écart-type	degrès de liberté	<i>F</i>	<i>p</i>
jardin	fort	41	24.02	3.62	118	21.19	0.0001
	moyen	46	20.63	4.37			
	faible	32	17.59	4.28			
1 ^{ère}	fort	61	26.57	3.92	134	6.69	0.0017
	moyen	68	24.16	4.12			
	faible	6	22.33	7.55			
2 ^e	fort	69	30.26	3.15	147	6.77	0.0016
	moyen	52	28.15	4.36			
	faible	27	27.74	3.83			
3 ^e	fort	67	31.96	2.96	167	12.09	0.0001
	moyen	81	30.10	3.71			
	faible	20	27.70	4.84			
4 ^e	fort	60	33.10	2.86	126	9.10	0.0002
	moyen	49	31.74	3.73			
	faible	18	28.94	5.49			
5 ^e	fort	56	34.34	2.24	138	13.24	0.0001
	moyen	70	32.30	3.34			
	faible	13	30.46	2.50			
6 ^e	fort	83	35.06	2.47	140	8.05	0.0005
	moyen	45	33.80	2.93			
	faible	13	32.31	2.25			

Les tests post-hoc de Scheffé révèlent qu'entre le groupe au rendement académique fort et celui au rendement académique faible, les différences sont significatives à tous les niveaux scolaires sauf en première année. Dans ce cas particulier, on ne peut pas procéder à l'interprétation des résultats de l'analyse statistique à cause du nombre très restreint d'élèves formant le groupe faible ($n=6$). Par contre, pour la compréhension orale, le test se révèle très sensible pour faire des distinctions non seulement entre les élèves forts et faibles mais également entre élèves forts et moyens puisque les différences sont significatives à tous les niveaux scolaires (sauf en quatrième année). Cette relation très marquée entre les résultats au test et le rendement académique en compréhension orale est particulièrement importante puisqu'il s'agit d'un test de compréhension syntaxique.

Des analyses semblables ont été effectuées pour le rendement académique en expression orale, mathématiques, lecture et rédaction. Les analyses de la variance montrent qu'entre les

Tableau 7. Résultats des tests post-hoc Scheffé en compréhension orale.

niveau scolaire	rendement académique	différence	Scheffé
jardin	fort vs moyen	3.39	7.03*
	fort vs faible	6.43	20.93*
	moyen vs faible	3.04	4.90*
1 ^{ère} année	fort vs moyen	2.41	5.27*
	fort vs faible	4.24	2.77
	moyen vs faible	1.83	0.52
2 ^e année	fort vs moyen	2.11	4.71*
	fort vs faible	2.52	4.41*
	moyen vs faible	0.41	0.11
3 ^e année	fort vs moyen	1.86	4.91*
	fort vs faible	4.26	10.83*
	moyen vs faible	2.40	3.58*
4 ^e année	fort vs moyen	1.36	1.87
	fort vs faible	4.16	8.91*
	moyen vs faible	2.79	3.82*
5 ^e année	fort vs moyen	2.04	7.85*
	fort vs faible	3.88	9.63*
	moyen vs faible	1.84	2.25
6 ^e année	fort vs moyen	1.26	3.41*
	fort vs faible	2.75	6.27*
	moyen vs faible	1.49	1.66

* $p < 0.05$

groupes fort et faible les différences sont significatives à tous les niveaux, dans toutes les catégories académiques (sauf en deuxième année, en lecture et en rédaction). Il est possible que ce niveau scolaire coïncide avec une phase de croissance particulièrement importante au niveau de la compréhension orale mais que le transfert ne soit pas encore fait au niveau du langage écrit. Le fait que les élèves soient forts ou faibles au niveau de la lecture et de la rédaction ne se traduirait pas par une différence significative dans un test de compréhension syntaxique puisque, à ce stade de développement du langage, le versant réceptif serait plus avancé.

Hypothèse 5 : les élèves qui reçoivent de l'aide académique supplémentaire réussiront moins bien au test que les élèves qui ne reçoivent pas cette aide

Le Tableau 8 présente les moyennes des élèves recevant de l'aide supplémentaire en comparaison à celles des élèves qui n'en reçoivent pas. Il est à noter qu'il existe des différences significatives à tous les niveaux scolaires sauf en deuxième année.

Tableau 8. Statistiques descriptives des résultats des élèves recevant de l'aide académique supplémentaire comparés à ceux qui n'en reçoivent pas et les ANOVAs.

niveau scolaire et aide (o/n)	n	moyenne des réponses correctes	écart - type	degrés de liberté	F	p
1 ^{ère}	oui	24	22.50	4.04	24	11.67
	non	111	25.75	4.26		
2 ^e	oui	24	27.92	4.78	24	2.51
	non	124	29.28	3.67		
3 ^e	oui	22	27.05	4.87	22	24.30
	non	146	31.08	3.35		
4 ^e	oui	18	27.94	5.50	18	27.33
	non	109	32.65	3.12		
5 ^e	oui	19	30.32	3.92	19	17.64
	non	120	33.37	2.77		
6 ^e	oui	14	32.43	2.28	14	8.53
	non	126	34.63	2.71		

Hypothèse 6 : les élèves ayant des difficultés de langage réceptif auront des scores moins élevés que les élèves sans difficultés de langage

Bien que dans les quatre écoles choisies il y ait eu 32 élèves déjà identifiés par une orthophoniste comme ayant des difficultés de langage, seuls 23 d'entre eux (soit 72%) ont participé à l'étude. Il est possible que les parents n'aient pas accordé leur autorisation soit parce que l'élève était déjà souvent retiré de la salle de classe, soit parce que l'élève avait eu de nombreuses évaluations. Ce taux de participation est différent de celui de l'échantillon complet qui s'élève à 90%.

Une analyse des 23 cas avec difficultés de langage montre que 14 de ces élèves (61%) ont des scores inférieurs au 15^e rang centile, ce qui est concordant avec la présence d'un problème significatif au niveau du langage. Trois élèves (13%) ont eu des scores compris entre le 15^e et le 30^e rang centile. Les six derniers cas (26%) qui ont réussi au-delà du 30^e rang centile ont été discutés avec l'orthophoniste impliquée dans chaque cas. Afin d'expliquer ces scores élevés, il a été possible pour trois élèves d'émettre les hypothèses suivantes. Pour un élève, les difficultés de langage se situaient surtout au niveau expressif. Le

Tableau 9. Rangs centile des scores bruts selon les normes originales anglaises et celles développées lors de l'étude selon l'âge.

groupe d'âge	score brut	rang centile au test anglais	rang centile au test français	
			indice de francité	
			faible	élevé
5;0 - 5;5	19	54	31	13
	13	26	8	8
	8	11	<1	<1
5;6 - 5;11	20	52	50	27
	15	26	15	<1
	9	10	<1	<1
6;0 - 6;11	25	48	47	38
	20	27	16	4
	17	11	3	<1
7;0 - 7;11	33	57	88	70
	28	24	38	16
	21	10	4	4
8;0 - 8;11	35	55	92	89
	31	24	55	51
	28	11	24	24
9;0 - 9;11	36	52	91	87
	33	23	67	55
	30	11	33	27

score élevé au test d'un deuxième élève pourrait être dû au phénomène d'apprentissage au test, cet élève ayant subi le même test à deux reprises dans les deux mois précédant sa participation au projet. Dans le cas d'un troisième élève, un problème d'hyperactivité a pu influencer les résultats de l'évaluation orthophonique.

Hypothèse 7 : les normes développées lors de ce projet seront différentes des normes du test original

Parmi les problèmes associés à l'utilisation de tests figure celui de l'usage incorrect des normes. Par exemple, l'utilisation des normes d'un test après avoir modifié l'outil de façon significative peut donner lieu à des erreurs d'interprétation. La traduction d'un test représente un changement suffisamment important pour qu'il soit nécessaire d'élaborer des normes propres au test modifié.

Le Tableau 9 permet de visualiser les différences qui existent entre les normes originales et celles obtenues lors de notre recherche. D'un point de vue clinique, un élève de six ans ayant un score brut de 20, obtiendrait un résultat dans la moyenne (27^{ème} rang centile) selon les normes originales anglaises. Les

résultats de ce même élève, calculés selon les normes développées lors de ce projet, ne se situeraient pas dans la moyenne (16^{ème} rang centile avec un indice de francité faible et 4^{ème} rang centile avec un indice de francité élevé).

Hypothèse 8 : l'ordre de difficulté des items sera semblable à celui retrouvé dans le test original

La section III *Elaborated Sentences* comporte quarante phrases classées par ordre croissant de difficulté grammaticale. Ces quarante phrases stimuli peuvent être également regroupées en huit catégories grammaticales selon le modèle de Carrow-Woolfolk (1985) qui a basé cette classification d'items d'après des recherches théoriques et des données empiriques. D'après son modèle, un item a été considéré comme étant acquis dès que 90% des élèves d'une même tranche d'âge réussissait cet item. Nous avons suivi le même exemple que Carrow-Woolfolk avec nos deux groupes ayant un indice de francité élevé (voir Figure 1) et faible et nous les avons comparés, catégorie par catégorie. Quoique certains items aient été réussis plus tôt par le groupe ayant un indice de francité élevé, l'ordre croissant des difficultés des items par catégorie a été semblable pour les deux groupes. Nous avons utilisé ensuite les résultats obtenus par le groupe ayant un indice de francité élevé et les avons comparés avec les résultats de Carrow-Woolfolk (1985, p. 92), catégorie par catégorie.

De façon globale, l'ordre croissant de difficulté par catégorie grammaticale est comparable. Néanmoins en français, à l'intérieur de certaines catégories, on trouve des items qui ne respectent pas l'ordre de difficulté. Les différences retrouvées dans l'ordre de difficulté pourraient être attribuables aux modifications qui ont dû être opérées sur les items lors de la traduction.

Hypothèse 9 : les validités de construit, de contenu et celle liée à un critère seront établies

Validité de construit. La validité de construit repose sur deux postulats :

1. L'élève doit maîtriser les structures syntaxiques simples avant de maîtriser les structures plus élaborées.
2. Plus l'élève est âgé, plus le nombre de structures syntaxiques maîtrisées est élevé.

Pour établir la validité de construit de la traduction de la section III du *TACL-R*, il fallait donc mesurer le degré d'adéquation de ces postulats avec les résultats au test.

Une analyse de la Figure 1 montre que les élèves maîtrisent des formes syntaxiques telles que les phrases actives avant de maîtriser des structures plus élaborées telles que les phrases avec enchâssement. Ceci montre que les résultats au test sont en accord avec le postulat affirmant que les structures simples sont maîtrisées avant les structures complexes.

En regardant les résultats des élèves par âge (Tableau 2), on voit que les scores des élèves augmentent en fonction de l'âge.

Une ANOVA a démontré que la différence est significative [$F(7,969) = 162.38; p < .0001$]. Le deuxième postulat est ainsi confirmé.

Validité de contenu. Comme pour tout test, il était nécessaire de vérifier la validité de contenu de la traduction/adaptation de la section III du *TACL-R*.

Le *TACL-R* est un test de compréhension syntaxique et son auteure, Elizabeth Carrow-Woolfolk, a établi la validité des items (Carrow-Woolfolk, 1985). En ce qui concerne la traduction française, une analyse de contenu des 39 items a été faite par les trois orthophonistes participant au projet. Ces personnes ayant une expertise en linguistique et des connaissances cliniques dans le domaine du langage étaient à même de déterminer la validité du contenu puisque d'après Gay, "content validity is

determined by expert judgement. There is no formula for computing it and there is no way to express it quantitatively" (1985, p. 160).

Tout comme dans la version anglaise, les items illustrent un éventail des structures syntaxiques de la langue. On retrouve des exemples de phrases interrogatives, négatives, à la voix active et passive, des propositions subordonnées et coordonnées et des phrases avec enchâssement. D'après leur expérience clinique et leur jugement professionnel, les orthophonistes ont été d'accord pour affirmer que le vocabulaire utilisé était simple et que la difficulté à résoudre dans l'item portait donc sur la syntaxe et non pas le vocabulaire.

Validité liée à un critère. Carrow-Woolfolk a comparé le *TACL-R* au *TACL* et à une variété d'autres tests. Les coefficients de corrélation qu'elle a trouvés sont tous acceptables. Ceci est encourageant pour la validité critérielle de la version française de la section III mais ne peut servir comme preuve de validité puisqu'une traduction peut changer de façon significative le contenu d'une épreuve. De plus, la traduction porte uniquement sur une partie du test original.

La validité critérielle du test a été examinée selon les critères suivants : rendement académique, aide académique supplémentaire et présence d'un problème de langage. Puisqu'il n'existe pas d'outils standardisés pour évaluer les habiletés en langage réceptif pour la population cible de cette étude, il a été décidé de comparer les résultats au test et le rendement académique des élèves dans les domaines suivants : compréhension orale, expression orale, mathématiques, lecture et rédaction.

Plusieurs études récentes indiquent que les estimations de rendement académique faites par les enseignants et les enseignantes sont habituellement plus ou moins justes, mais en général elles ne se révèlent pas fausses (Egan & Archer, 1985; Gerber & Semmell, 1984; Gresham, Reschly, & Carey, 1987; Hoge, 1983).

La validité critérielle basée sur le rendement académique a été établie lors de la discussion de l'hypothèse 4.

Le Tableau 8 montre l'existence de différences significatives entre les résultats obtenus par les élèves recevant de l'aide académique supplémentaire et ceux n'en recevant pas. Ces résultats sont une autre

Figure 1 : Classification des items par catégorie grammaticale selon l'âge de réussite, pour les élèves avant l'indice de francité élevé.

Questions	5;0	5;11	6;0	6;11	7;0	7;11	8;0	8;11	9;0 - 9;11	10;0 - 10;11	11;0 - 12;09	Indéterminé
interrogatives												
1												
13												
23												
25												
30												
négatives												
4												
5												
8												
11												
actives												
6												
14												
passive												
16												
C.O.D./C.O.I												
17												
28												
coordonnées												
2												
3												
7												
9												
26												
29												
36												
subordonnées												
10												
18												
20												
21												
27												
31												
32												
33												
37												
enchâssement												
12												
15												
19												
22												
34												
35												
38												
39												
40												

Un item est acquis dès que 90 % des élèves d'une même tranche d'âge réussissent cet item.

preuve de la validité critérielle du test. Si un élève reçoit de l'aide supplémentaire c'est qu'il est considéré par les enseignants et les enseignantes comme ayant des problèmes importants et le test a permis d'identifier ces élèves.

Tel que discuté dans l'hypothèse 6, la validité critérielle du test a également été établie d'après le critère de la présence d'un problème de langage.

Hypothèse 10 : la fidélité du test sera démontrée

Une méthode pour déterminer la fidélité d'un test consiste à calculer l'erreur-type de mesure. Les Tableaux 2 et 5 montrent que les valeurs des erreurs-types de mesure pour chaque groupe d'âge se situent entre 0.25 et 0.83.

Une autre méthode est celle du test répété. Puisque seulement 5% des élèves ont passé le test à deux reprises, il n'a pas été possible d'effectuer une analyse statistique par groupe d'âge. Une analyse globale des résultats met en évidence une corrélation élevée ($r=.869$, $n=65$, $p < .05$). Dans la plupart des cas les scores au test sont semblables, l'écart allant de 0 à 4 points. Pour quelques élèves, la différence était supérieure à cinq points; la plupart de ces élèves étaient parmi les plus jeunes.

Une troisième méthode pour mesurer la fidélité d'un test est de calculer le coefficient de consistance interne. La consistance interne détermine si le test est homogène. Le test a été divisé en deux en fonction des questions paires et impaires. La formule de nivellement de Spearman-Brown a été employée pour ajuster les coefficients de corrélation selon la longueur du test. Les coefficients de corrélation (le r de Bravais-Pearson) pour chaque groupe d'âge se situent entre 0.6 et 0.7 indiquant une corrélation modérée entre les deux moitiés du test. Par contre, le coefficient pour l'échantillon complet est élevé avec une valeur de 0.824.

Conclusion

Le but de la recherche présentée dans le cadre de cet article consistait à étalonner un instrument d'évaluation en compréhension syntaxique d'élèves fréquentant les écoles de langue française du centre-sud de l'Ontario.

Cette étude révèle que les résultats obtenus aux niveaux de la validité et de la fidélité de l'adaptation de la section III du TACL-R sont acceptables. Les tests de fidélité ont révélé des coefficients modérés et élevés. L'analyse qualitative et quantitative de la validité externe et interne indique qu'en général le test est valide. Afin d'obtenir des normes pour notre population cible, nous nous sommes servies d'une approche unique qui tenait compte de l'indice de francité de cette population. Les normes régionales obtenues sont différentes des normes originales, ce qui confirme qu'il est tout à fait contre-indiqué d'utiliser les normes originales d'un test auprès d'une population au vécu linguistique différent.

Cette recherche est une étape importante vers une meilleure compréhension des compétences linguistiques de notre population. Nous espérons que ces résultats encourageront d'autres chercheurs à explorer les différentes facettes des habiletés langagières des populations vivant en milieu linguistique minoritaire car ces études futures permettraient de mieux saisir les mécanismes du langage dans ces situations particulières.

Prière d'envoyer toute correspondance à : Lynne Latulippe, Conseil des écoles séparées catholiques romaines de Dufferin et Peel, 40, boul. Matheson ouest, Mississauga, Ontario L5R 1C5.

Note de la rédaction : Sylvie Gautheron-Stone était anciennement au Conseil des écoles françaises de la communauté urbaine de Toronto et au Conseil des écoles séparées catholiques romaines de Hamilton-Wentworth. Linda Siegel était anciennement au Département de psychologie appliquée de l'Institut d'études pédagogiques de l'Ontario.

Remerciements

Nous adressons toute notre reconnaissance aux Fonds de langue française du Ministère de l'Éducation et de la Formation de l'Ontario pour son appui financier. Nous exprimons notre gratitude aux quatre conseils sans la participation desquels ce projet n'aurait pu aboutir. Nous tenons à remercier tout spécialement nos deux agentes de recherche, Sylvie Lemay et Marie-Denise Bélec qui ont collaboré à la collecte des données et à l'analyse statistique des résultats et dont l'enthousiasme a été un atout au cours de cette recherche. Enfin, nous saluons la contribution précieuse de deux orthophonistes, Nicole Moore et Marie-Noël Olivier.

Références

- Bankson, N. (1989). Review of the Test for Auditory Comprehension of Language – Revised. In J. C. Conoley & J. J. Kramer (Eds.), *The tenth mental measurements yearbook* (pp. 822-824). Lincoln, NB : University of Nebraska Press.
- Carrow, E. (1973). *Test for Auditory Comprehension of Language*. Allen, TX : DLM Teaching Resources.
- Carrow-Woolfolk, E. (1985). *Test for Auditory Comprehension of Language – Revised*. Allen, TX : DLM Teaching Resources.
- Desjarlais, L. (1980). *L'influence du milieu socio-linguistique sur les élèves franco-ontariens : une étude de cas*. Toronto : Ministère de l'Éducation.
- Dudley, J., & Delage, J. (1980). *Tests de langage dudley/delage*. Saint Lambert, QC : Les Éditions de l'ABC.
- Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test – Revised*. Circle Pines, MI : American Guidance Service.

- Dunn, L. M., Thériault-Whalen, C., & Dunn, L. M. (1993). *L'Échelle de vocabulaire en images Peabody*. Toronto : Psycan.
- Egan, O., & Archer, P. (1985). The accuracy of teachers ratings of ability: A regression model. *American Educational Research Journal*, 22, 25-34.
- Gay, L. R. (1985). *Educational evaluation and measurement*. Columbus, OH : Charles Merrill.
- Gerber, M., & Semmel, M. (1984). Teachers as imperfect tests: Reconceptualizing the referral process. *Educational Psychologists*, 19, 137-148.
- Gresham, F. M., Reschly, D. J., & Carey, M. P. (1987). Teachers as "tests": Classification accuracy and concurrent validation in the identification of learning disabled children. *School Psychology Review*, 16, 543-553.
- Haynes, W. (1989). Review of the Test for Auditory Comprehension of Language - Revised. In J. C. Conoley & J. J. Kramer (Eds.), *The tenth mental measurements yearbook* (pp. 824-826). Lincoln, NB : University of Nebraska Press.
- Hébrard, P., & Mougeon, R. (1975). La langue parlée le plus souvent entre les parents et les enfants : un facteur crucial dans l'acquisition linguistique de l'enfant dans un milieu bilingue. *Working Papers in Bilingualism*, 7, 54-70.
- Heller, M., & Barker, G. (1988). Conversational strategies and contexts for talks, learning activities for franco-ontarian minority schools. *Anthropology & Education Quarterly*, 19, 20-47.
- Heller, M., Levy, L., Mougeon, R., & Quirouette, P. (1986). La migration et l'immigration francophone en Ontario : leur incidence sur l'éducation, l'assimilation et la mobilité sociale. Document préparé pour l'étude de la démographie et son incidence sur la politique sociale et économique. Toronto : Centre de recherches en éducation franco-ontarienne (CRÉFO).
- Hoge, R. D. (1983). Psychometric properties of teacher-judgement measures of pupil aptitudes, classroom behaviors and achievement levels. *The Journal of Special Education*, 17, 401-429.
- Latulippe, L. (1992). Réflexions d'une orthophoniste oeuvrant en milieu linguistique minoritaire. *Communiqué*, 6 (4), 15-17.
- Mougeon, R. (1979). Les mariages mixtes et l'assimilation des francophones. *Revue de l'association canadienne d'éducation de langue française*, 8, 24-27.
- Mougeon, R., & Heller, M. (1986). The social and historical context of minority French language education in Ontario. *Journal of Multilingual and Multicultural Development*, 7, 199-227.
- Mougeon, R., Heller, M., Béniak, E., & Canale, M. (1984). Acquisition et enseignement du français en situation minoritaire : le cas des francoontariens. *La revue canadienne des langues vivantes*, 41, 315-335.
- Rondal, J. A., & Seron, X. (1985). *Troubles du langage : diagnostic et rééducation*, Bruxelles : Pierre Mardaga.
- Roussel, N. (1991). Annotated bibliography of communicative ability tests. In E. V. Hamayan & J. S. Damico (Eds.), *Limiting bias in the assessment of bilingual students* (p. 230-343). Austin, TX : Pro-Ed.
- Statistique Canada. (1992). *Langue maternelle, Recensement du Canada de 1991*, Ottawa, ON : Approvisionnement et Services Canada.
- Wilcox, K., & McGuinn-Aasby, S. (1988). The performance of monolingual and bilingual Mexican children on the TAFL. *Language, Speech and Hearing Services in Schools*, 19, 34-40.